

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/777,010A  
Source: FWO  
Date Processed by STIC: 11/8/06

# ***ENTERED***



IFWO

## RAW SEQUENCE LISTING

DATE: 11/08/2006

PATENT APPLICATION: US/10/777,010A

TIME: 09:40:20

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\11082006\J777010A.raw

3 <110> APPLICANT: Carstens, Carsten P.  
 5 <120> TITLE OF INVENTION: Methods and Compositions for High Level Expression of a  
 6 Heterologous Protein with Poor Codon Usage  
 8 <130> FILE REFERENCE: 225436/1344  
 10 <140> CURRENT APPLICATION NUMBER: 10/777,010A  
 11 <141> CURRENT FILING DATE: 2004-02-11  
 13 <150> PRIOR APPLICATION NUMBER: US 09/492,590  
 14 <151> PRIOR FILING DATE: 2000-01-27  
 16 <150> PRIOR APPLICATION NUMBER: US 60/117,355  
 17 <151> PRIOR FILING DATE: 1999-01-27  
 19 <160> NUMBER OF SEQ ID NOS: 16  
 21 <170> SOFTWARE: PatentIn version 3.3  
 23 <210> SEQ ID NO: 1  
 24 <211> LENGTH: 29  
 25 <212> TYPE: DNA  
 26 <213> ORGANISM: Artificial sequence  
 28 <220> FEATURE:  
 29 <223> OTHER INFORMATION: PCR Primer for E. coli argU gene  
 31 <400> SEQUENCE: 1  
 32 gacactagta atcagacgcg gtcgttcac 29  
 35 <210> SEQ ID NO: 2  
 36 <211> LENGTH: 32  
 37 <212> TYPE: DNA  
 38 <213> ORGANISM: Artificial sequence  
 40 <220> FEATURE:  
 41 <223> OTHER INFORMATION: PCR primer for E. coli ArgU gene  
 43 <400> SEQUENCE: 2  
 44 gacgacgaca agaatcagac gcggtcgttc ac 32  
 47 <210> SEQ ID NO: 3  
 48 <211> LENGTH: 29  
 49 <212> TYPE: DNA  
 50 <213> ORGANISM: Artificial sequence  
 52 <220> FEATURE:  
 53 <223> OTHER INFORMATION: PCR primer for E. coli ArgU gene  
 55 <400> SEQUENCE: 3  
 56 ctgccatggt ggaggatata aagaaggcg 29  
 59 <210> SEQ ID NO: 4  
 60 <211> LENGTH: 29  
 61 <212> TYPE: DNA  
 62 <213> ORGANISM: Artificial sequence  
 64 <220> FEATURE:  
 65 <223> OTHER INFORMATION: PCR primer for E. coli IleY gene  
 67 <400> SEQUENCE: 4

## RAW SEQUENCE LISTING

DATE: 11/08/2006

PATENT APPLICATION: US/10/777,010A

TIME: 09:40:20

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\11082006\J777010A.raw

```

68 cagccatggc cttgaaatgg cgtagtca                29
71 <210> SEQ ID NO: 5
72 <211> LENGTH: 29
73 <212> TYPE: DNA
74 <213> ORGANISM: Artificial sequence
76 <220> FEATURE:
77 <223> OTHER INFORMATION: PCR primer for E. coli IleY gene
79 <400> SEQUENCE: 5
80 gacactagtc cttgaaatgg cgtagtca                29
83 <210> SEQ ID NO: 6
84 <211> LENGTH: 29
85 <212> TYPE: DNA
86 <213> ORGANISM: Artificial sequence
88 <220> FEATURE:
89 <223> OTHER INFORMATION: PCR primer for E. coli IleY gene
91 <400> SEQUENCE: 6
92 cagtctagat catcatgttt attgcgtgg                29
95 <210> SEQ ID NO: 7
96 <211> LENGTH: 29
97 <212> TYPE: DNA
98 <213> ORGANISM: Artificial sequence
100 <220> FEATURE:
101 <223> OTHER INFORMATION: PCR primer for E. coli IleY gene
103 <400> SEQUENCE: 7
104 gacctcgagt catcatgttt attgcgtgg                29
107 <210> SEQ ID NO: 8
108 <211> LENGTH: 29
109 <212> TYPE: DNA
110 <213> ORGANISM: Artificial sequence
112 <220> FEATURE:
113 <223> OTHER INFORMATION: PCR primer for E. coli LeuW gene
115 <400> SEQUENCE: 8
116 cagtctagag aatcccgtcg tagccacca                29
119 <210> SEQ ID NO: 9
120 <211> LENGTH: 30
121 <212> TYPE: DNA
122 <213> ORGANISM: Artificial sequence
124 <220> FEATURE:
125 <223> OTHER INFORMATION: PCR primer for E. coli LeuW gene
127 <400> SEQUENCE: 9
128 gacctcgagg gcatccgatc aacgctttct                30
131 <210> SEQ ID NO: 10
132 <211> LENGTH: 29
133 <212> TYPE: DNA
134 <213> ORGANISM: Artificial sequence
136 <220> FEATURE:
137 <223> OTHER INFORMATION: PCR primer for E. coli ProL gene
139 <400> SEQUENCE: 10
140 gacgtcgacg tgctgacaga cgagaagcg                29

```

## RAW SEQUENCE LISTING

DATE: 11/08/2006

PATENT APPLICATION: US/10/777,010A

TIME: 09:40:20

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\11082006\J777010A.raw

```

143 <210> SEQ ID NO: 11
144 <211> LENGTH: 29
145 <212> TYPE: DNA
146 <213> ORGANISM: Artificial sequence
148 <220> FEATURE:
149 <223> OTHER INFORMATION: PCR primer for E. coli ProL gene
151 <400> SEQUENCE: 11
152 gacctcgagg gtgtggtctg gacgttctg                29
155 <210> SEQ ID NO: 12
156 <211> LENGTH: 29
157 <212> TYPE: DNA
158 <213> ORGANISM: Artificial sequence
160 <220> FEATURE:
161 <223> OTHER INFORMATION: PCR primer for E. coli GlyU gene
163 <400> SEQUENCE: 12
164 ctgccatggg gcacttgcta aggagagcg                29
167 <210> SEQ ID NO: 13
168 <211> LENGTH: 33
169 <212> TYPE: DNA
170 <213> ORGANISM: Artificial sequence
172 <220> FEATURE:
173 <223> OTHER INFORMATION: PCR primer for E. coli GlyU gene
175 <400> SEQUENCE: 13
176 ggaacaagag ggcgtgtttt cctgggttgt tac                33
179 <210> SEQ ID NO: 14
180 <211> LENGTH: 10133
181 <212> TYPE: DNA
182 <213> ORGANISM: Escherichia coli
184 <400> SEQUENCE: 14
185 atgctatcag catggatgaa cggggcgtag agggcaaaag tctgaaaaga gaaccggcct                60
187 gttgatacag gccgggaaag ggatcaggca acaacctgta cgctgtgacc tgcaaaactc                120
189 actgtctgac cggcgacgat tttgcagcgt ttgcgcgttt caaccgcacc gtcgactttc                180
191 acctggcctt cggcaatcgc gatttttcgcc tgcgcgcgcg tttcgctcca gccttccagt                240
193 ttcagcaagt cgcacagctc aacgtgcgga tgtttaccta aagaaaatgt cgccatgtta                300
195 ctcatcctgt ggatcatgat attcaacgca cgctctgtagc gtgttttcaa tcagcgtggc                360
197 aaccgtcatc gggccaacgc cgccgggaac aggcgtaatg tatgaggcgc gtttagccgc                420
199 gtcttcaaac acgacgtcgc ccacaacttt gccattttcc agacggttga tgccgacatc                480
201 aatcacaatt gcgccttctt tgatccagtc accgggaata aagcctggct tgccaacggc                540
203 aacgatcaat agatcggcat tttctacgtg atgacgcaga tttttagtga agcggtgagt                600
205 cactgtagtg gtgcaacctg ccagcagcag ttccatgctc atcgggcggc caacgatatt                660
207 cgatgcgcca atcaccacgg cgttgaggcc gaaggtatca atgttgtaac gctcaagcag                720
209 cgtgacgata ccgcgcgggg tgcagggacg cagacgcggc gcgcgctggc acagacgacc                780
211 gacgttgtaa ggatggaaac cgtccacgct tttgtccgga tgaatacgtt ccagcacttt                840
213 gacgttatca ataccgccg gtaacggcag ttgaaccaga atgccatcga tgggtgtgtc                900
215 ggcattcagc gtatcgataa gctccagcag ctccgcttcg ctgggtggtt ccgggaggtc                960
217 ataagagcgg gagacgaacc cgacttcttc acaagccttg cgtttgcttg cgacataaat                1020
219 ttgcgatgca gggttactac ccaccagcac aacggccagt cctgggtgcc gcagtcgggc                1080
221 tgcaatacgc gcctgaactt tttgagcaac ttcagagcgc acctgctgcg caatcgtttt                1140
223 accgtcaata atctttgctg ccatcagaga gaggattcca tctgttacgt agatcgaagg                1200

```

## RAW SEQUENCE LISTING

DATE: 11/08/2006

PATENT APPLICATION: US/10/777,010A

TIME: 09:40:20

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\11082006\J777010A.raw

225	ggatgcgcct	attttgtcag	aagcggggcg	cgctgtcagg	tttcgtttca	gattttatcgc	1260
227	gtgaagcgac	ctcttgcgaa	ggtgaggcgc	accgtcgctg	agactgaaag	cttcattttt	1320
229	cgtccatgat	ggcgttgtaa	atctggaact	gattttatttc	cttgtctaag	gattaagata	1380
231	atttaagaaa	tacctgacaa	tataaaaaga	attttcagcc	tggtaattha	ccgcttcagg	1440
233	tctatatttg	tgttgaatat	attttgcgcg	gaagtattca	tctaacgggg	ctctctatth	1500
235	tttagaatag	agtgcataat	ttcaattaag	acattcttag	aggataaaaa	ggaatttact	1560
237	actatcagtg	tcttaaataa	agtaatcggg	tatatacgga	tgtggagtcg	ataaatgaga	1620
239	ttgaaggaat	atatatgaaa	ttaagattta	tttcgtctgc	gctggctgcc	gcactattcg	1680
241	ccgctacggg	tagttatgct	gccgttgtag	atggcggtag	aattcacttt	gaaggcgaac	1740
243	tggtgaatgc	tgctgttcca	gtgaataact	actcggcaga	ccaggttgtc	acactcggtc	1800
245	aatatcgtac	cgatatcttc	aatgctgttg	gtaatacctc	tgcattaatt	ccattcacca	1860
247	ttcagttgaa	cgactgcat	cctgttggtg	ccgctaattgc	tgccggttgc	ttttctggtc	1920
249	aggetgatgc	aatcaatgat	aattttattgg	ccattgcata	cagtaccaat	acaacaacag	1980
251	caacgggtgt	cggatttgaa	atacttgata	atacatccgc	aattctcaaa	cctgatggga	2040
253	atagcttctc	aaccaaccag	aacttgatcc	ccgggaccaa	cgttcttcat	ttttctgcac	2100
255	gttataaagg	caccggtaca	agtgcatacag	cagggaagc	aaatgctgac	gcgaacttta	2160
257	ttatgagata	tgaataatca	aaaccacggt	gttttgaatt	atatatcacg	tcttataaca	2220
259	aagtaatgta	ccggttgtct	gaagcgggat	ggtggcaatg	taaatcgaaa	tcatgttcac	2280
261	tttgtatcat	gccgctttat	taaatgaaaa	gggaatgatg	tgttgtaaga	aaccaaagca	2340
263	atcattttct	tatattcctt	atttttgccc	tcaggaatac	acaaggcgta	ttaactatga	2400
265	tgactaaaat	aaagttattg	atgctcatta	tattttatth	aatcatttcg	gccagcgccc	2460
267	atgctgcccg	agggatcgca	ttaggtgcca	cgcgtattat	ttatcccgt	gatgctaaac	2520
269	agactgcggg	atggattaga	aatagccata	ccaatgagcg	ctttctggtc	aattcgtgga	2580
271	ttgaaaacag	cagcgggtga	aaagaaaagt	cattcatcat	tacaccgcca	ctgtttgtta	2640
273	gtgaacccaa	aagcgaataa	actttgcgta	ttattttacac	cgggtccaccg	ctggcagcag	2700
275	atcgtgagtc	tctgttctgg	atgaatgtta	agacgatccc	ttcggtagat	aaaaatgcct	2760
277	tgaacggcag	gaatgttttg	caactggcga	ttttatcgcg	catgaaatta	tttctccgtc	2820
279	caattcaatt	acaagaatta	cccgcagaag	cgcgggacac	actcaagttt	tcgcgatccg	2880
281	gtaactatat	caatgttcat	aatccatcac	ctttttatgt	caccctgggt	aacttacaag	2940
283	tgggcagcca	aaagttgggg	aatgctatgg	ctgcacccag	agttaattca	caaattccct	3000
285	taccctcagg	agtgcaggga	aagctgaaat	ttcagaccgt	taatgattat	ggttcagtaa	3060
287	ctccggtcag	agaagtgaac	ttaaactaac	cgaatcatct	gacaatatca	gagctaatta	3120
289	tgaaaatacc	cactactacg	gatattccgc	agaggtatac	ctgggtgtctg	gccggaattt	3180
291	gttattcatc	tcttgccatt	ttaccctcct	ttttaagcta	tgcggaaagt	tatttcaacc	3240
293	cggcattttt	attagagaat	ggcacatccg	ttgctgattt	atcgcgcttt	gagagaggta	3300
295	atcatcaacc	tgccggcggtg	tatcggggtg	atctctggcg	taatgatgag	ttcattgggt	3360
297	cgcaggatat	cgtatttgaa	tcgacaacag	aaaatacagg	tgataaatca	ggtgggttaa	3420
299	tgcctgtttt	taaccaggta	cttcttgaa	gaattggcct	taatagcagt	gcatttcccg	3480
301	agttagccca	gcagcaaaac	aataaatgca	tcaatttact	gaaagctgta	cctgatgcca	3540
303	caattaactt	tgatttttga	gcgatgcgcc	tgaacatcac	tattcctcag	atagcgttgt	3600
305	tgagtagcgc	tcacgggttac	attccgcctg	aagagtggga	tgaaggtatt	cctgctttac	3660
307	tcttgaatta	taatttcacc	ggtaacagag	gtaattggtta	cgatagctat	ttttttagt	3720
309	agctcagcgg	gatttaatat	ggcccgtggc	gtttacgcaa	caatgggtcc	tggaaactat	3780
311	ttcgcggaaa	tggatatcat	tcagaaacgt	ggaaataat	tggcacctgg	gtacagcgcg	3840
313	ccattattcc	gctgaaaagt	gaactggtaa	tgggagacgg	caatacagga	agtgatattt	3900
315	tcgatggcgt	tggatttcgt	ggtgtacggc	tttattcttc	tgataatatg	tatcctgata	3960
317	gccagcaagg	gtttgcccc	acggtacgtg	ggattgcccg	tacggcgggc	cagctaacga	4020
319	ttcggcaaaa	tgggttttatt	atctatcaaa	gctatgtttc	ccccggcgct	tttgaaatta	4080
321	cagatttgca	cccgcacatc	tcaaattggcg	atctggacgt	caccatcgac	gagcgcgatg	4140

## RAW SEQUENCE LISTING

DATE: 11/08/2006

PATENT APPLICATION: US/10/777,010A

TIME: 09:40:20

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\11082006\J777010A.raw

323	gcaatcagca	gaattacaca	attccgtatt	caacagtgcc	aattttacaa	cgcggaagggc	4200
325	gtttcaaatt	tgacctgacg	gcggggcgatt	ttcgtagcgg	taatagtcag	caatcatcgc	4260
327	ctttcttttt	tcaggggtacg	gcactcggcg	gtttaccaca	ggaatttact	gcctacggcg	4320
329	ggacgcaatt	atctgccaat	tacaccgcct	ttttattagg	gctggggcgc	aatctcggga	4380
331	actggggcgc	agtgtcgtcg	gatgtaacgc	atgcgcgcag	tcagttagcc	gacgccagtc	4440
333	gtcatgaggg	ggattctatt	cgcttcctct	atgcgaaatc	gatgaacacc	ttcggcacca	4500
335	attttcagtt	aatgggttac	cgctattcga	cacaagggtt	ttataccctt	gatgatgttg	4560
337	cgtatcgtcg	aatggagggg	tacgaatatg	attacgacgg	tgagcatcgc	gatgaaccga	4620
339	taatcgtgaa	ttaccacaat	ttacgcctta	gccgtaaaga	ccgtttgcag	ttaaatgttt	4680
341	cacaatcact	taatgacttt	ggctcgtctt	atatttctgg	tacctatcaa	aaatactgga	4740
343	atacttcgga	ttcagatacg	tggatatcagg	tggggtatac	cagcagctgg	gttggcatca	4800
345	gttattcgtc	ctcattttcg	tggaatgaat	ctgtagggat	ccccgataac	gaacgtattg	4860
347	tcggacttaa	tgtttcagtg	cctttcaatg	ttttgaccac	acgtcgtctc	accggggaaa	4920
349	atgcgctcga	ccgcgcttat	gcctccttta	acgccaacgc	taacagcaac	gggcaaaata	4980
351	gctggctggc	aggtgtaggt	gggaccttac	tgggaaggcca	caacctgagt	tatcacgtaa	5040
353	gccaggggtga	tacctcgaat	aatgggtaca	cgggcagcgc	cacggcaaac	tggcaggccg	5100
355	cttacgggtac	gctggggggc	gggtataact	acgaccgcga	tcaacatgac	gttaactggc	5160
357	agctgtctgg	cggtgtggtc	gggcatgaaa	atggcataac	gctgagccag	cctttagggg	5220
359	ataccaatgt	tttgattaaa	gcgcctggcg	caggcgggtg	acgcattgaa	aatcaaactg	5280
361	gcattttaac	cgactggcgc	ggctatgcgg	tgatgctgta	tgccacgggt	tatcgggtata	5340
363	accgtatcgc	gcttgatacc	aatacgtatg	ggaattccat	cgatgttgaa	aaaaatatta	5400
365	gcagcgttgt	gccgacgcaa	ggcgcgttgg	ttcgtgccaa	ttttgatacc	cgcataggcg	5460
367	tgcgggcgct	cattaccggt	acccagggcg	gaaaaccggg	gccgtttgga	tactgggtac	5520
369	gggaaaacag	taccggaata	accagtatgg	tgggtgatga	cgggcaagtt	tatttaagtg	5580
371	gtgcgccatt	gtctgggtgaa	ttactgggtc	agtggggaga	cggcgcgaac	tcacgctgca	5640
373	ttgcgcacta	tgtattgccg	aagcaaaagt	tacagcaagc	cgctactggt	atttcggcag	5700
375	tttgacacac	tcctggctca	taaaggaaat	tatcaataag	ataatctgca	gattattatt	5760
377	ggcgatggca	tgtttgtgtc	tggcaaacat	atcctgggct	actgtttgtg	caaatagtac	5820
379	tggcgtagca	gaagatgaac	actatgatct	ctcaaatact	tttaatagca	ccaataacca	5880
381	gccagggcag	attgttgttt	taccggaaaa	atccggctgg	gtaggtgtct	cagcaatttg	5940
383	tcaccccggt	acgctgggtga	attatacata	ccgtagtatt	gtcaccaact	ttattgttca	6000
385	ggaaactatc	gataattata	aatatatgca	attacatgat	tatctattag	gtgcgatgag	6060
387	tctggttgat	agtgtgatgg	atattcagtt	ccccccgcaa	aattatattc	ggatgggaac	6120
389	agatcctaac	gtttcgcaaa	accttccatt	cggggtgatg	gattctcgtt	taatatttcg	6180
391	tttaaagggt	attcgtccct	ttattaacat	ggtggagatc	cccagacagg	tgatgtttac	6240
393	cgtgtatgtg	acatcaacgc	cttacgatcc	gttgggtaca	cctgtttata	ccattagttt	6300
395	tgggtggccgg	gttgaagtac	cgcaaaaactg	cgaattaaat	gccgggcaga	ttgttgaatt	6360
397	tgatttttgt	gatatcggcg	catcgttatt	tagtgcgcca	gggcccgggt	atcgacctgc	6420
399	tgggtgtcatg	ccgcaaacca	agagcattgc	ggtcaaatgt	acgaatgttg	ctgcgcaggc	6480
401	ttatttaaca	atgcgtctgg	aagccagtgc	cgtttctggt	caggcgatgg	tgtcggacaa	6540
403	tcaggattta	ggttttattg	tcgccgatca	gaacgatacg	ccgatcacgc	ctaacgatct	6600
405	caatagcggt	attcctttcc	gtctggatgc	agctgcggca	gccaatgtca	cacttcgcgc	6660
407	ctggcctatc	agtattaccg	gtcaaaaacc	gaccgaaggg	ccgtttagcg	cgctggggta	6720
409	tttagcgcgc	gattatcaat	gaggtacgga	gaatgagaag	agtactcttt	agctgtttct	6780
411	gcgggctact	gtggagttcc	agtggtatgg	cagttgacct	tttaggaacg	attaatatca	6840
413	atttgcacgg	taacgttgtt	gattttctct	gtaccgtaaa	cacagcggat	attgataaga	6900
415	cggtagattt	aggcagatgg	cctacgacac	aactactgaa	cgctggcgat	accacggcac	6960
417	tcgtcccttt	tagcctgcgg	ctggagggat	gtcctccggg	ttcagttgcg	attttattta	7020
419	cggaacgcc	ggcatccgat	accaacctgc	tggctctgga	tgatcccgca	atggcacaaa	7080

**VERIFICATION SUMMARY**

DATE: 11/08/2006

PATENT APPLICATION: US/10/777,010A

TIME: 09:40:21

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\11082006\J777010A.raw